

Default probabilities, CDS premiums and downgrades : A probit-MIDAS analysis

Citation for published version (APA):

Freitag, L. (2014). *Default probabilities, CDS premiums and downgrades : A probit-MIDAS analysis*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 038 <https://doi.org/10.26481/umagsb.2014038>

Document status and date:

Published: 01/01/2014

DOI:

[10.26481/umagsb.2014038](https://doi.org/10.26481/umagsb.2014038)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Lennart Freitag

**Default probabilities, CDS
premiums and Downgrades:
A Probit-MIDAS Analysis**

RM/14/038

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O. Box 616
NL- 6200 MD Maastricht
The Netherlands

Default probabilities, CDS Premiums and Downgrades: A Probit-MIDAS Analysis

Lennart Freitag *

October 27, 2014

Abstract

This paper examines the relationship between sovereign credit default swaps (CDS) and sovereign rating changes of European countries. To this aim, a new estimator is introduced which merges mixed data sampling (MIDAS) with probit regression. Simulations show that the estimator has good properties in finite sample. Also, I investigate a bootstrap procedure introduced by Ghysels et al. (2007), which should be able to handle significance testing in a MIDAS setting. The bootstrap has good size but low power. For the empirical analysis I use sovereign CDS data for 22 EU countries trying to correlate sovereign downgrades with sovereign CDS premiums. Overall the CDS data and the ratings are in most cases significantly positively correlated. Therefore, Credit Rating Agencies (CRA) and financial markets are generally agreeing on the implied default probability of sovereign nations. Also, CDS prices are expecting downgrades in advance in the majority of investigated datasets. However, this does not mean that a default probability can be extracted from raw CDS prices. Instead, by using a MIDAS estimator, I significantly reduce the amount of noise in the data. Therefore, CRAs are still providing important information to financial markets.

JEL Classification: C25, C51, G24

Keywords: rating agency, sovereign credit rating, MIDAS, European Debt Crisis, CDS

*Department of Economics, Maastricht University, PO box 616, 6200 MD Maastricht, the Netherlands, l.freitag@maastrichtuniversity.nl

Lennart Freitag would like to thank Thomas Götz, Lenard Lieb, Alain Hecq, and Bertrand Candelon for feedback and discussion while writing this paper

1 Introduction

This paper investigates whether and how sovereign rating downgrades are correlated with credit default swaps premiums. A sovereign CDS is an insurance on the default of a government issued bond, while a sovereign rating gives the likelihood of the default of a country. Therefore, the premium of a CDS essentially measures the same as a rating. In this paper I want to investigate whether downgrades and CDS premiums are actually as strongly correlated with each other, as theory predicts. The main reason for only analyzing downgrades is that in the time-frame for which there are sufficient sovereign CDS premiums available (i.e. from 2006 onwards), we see nearly exclusively downgrades for European countries.

The initial seller of a CDS is a bank or insurance company. The buyer pays a fee to the seller, and in return receives the nominal value of the underlying asset, in case said asset defaults. This fee may be paid in installments (which is called a spread) or as a one time sum (which is called a premium). The fee is calculated using no-arbitrage arguments. The literature on the determinants of sovereign CDS prices is quickly growing. There are two main strands of the literature. The first one claims that most of the variation of sovereign CDS can be explained by global factors such as the state of US financial markets and its economy. Pan and Singleton (2008) find systemic risk in the credit-event arrival for three countries using a one-factor lognormal model. Ang and Longstaff (2013) use a multifactor affine model to disentangle systemic and local effects on CDS spreads. They come to the conclusion that macroeconomic variables are less important than the systemic risk created by global financial markets. The second strand of the literature claims that local risk factors are mainly driving CDS spreads. Alter and Schüller (2012) look at interdependencies between sovereign CDS and financial institution CDS using multi-equation regressions and conclude that the interaction between actors in the local financial industry is driving sovereign CDS spreads in the post financial crisis world. Several papers merge the two strands of literature. Using panel regression techniques, Arghyrou and Kontonikas (2012) find that it was only after 2007 that EMU CDS spreads are driven by macroeconomic fundamentals, while before 2007 EMU fundamentals were not significant. Remolona et al. (2008) decompose CDS spreads into expected losses and market risk premiums and find that the former is explained by macroeconomic variables, while the latter is explained by global risk factors.

Ratings have a very similar interpretation than CDS, however ratings are created in a completely different fashion. A rating agency is a corporation which is specialized in assessing the probability of default of entities (including but not limited to firms and states) or financial assets (such as collateralized debt obligations). This assessment is published via a rating which is similar to grades. Therefore, the probability is not explicitly communicated. The market for ratings has been dominated for decades by three big rating agencies, namely Moody's, Fitch, and Standard and Poor's (S & P).¹ Recently, these agencies have come under severe criticism for their role in the 2007/2008 financial crisis as well as their behavior in the European debt crisis.

Research about the determinants of sovereign credit ratings started in the nineties. Cantor and Packer (1996) are among the first to investigate them and established a set of macroeconomic variables that explain sovereign ratings. Afonso (2003) is refining this line of research. In his paper the author is proposing a list of variables that are significantly correlated with ratings as well as looking at different transformation of ratings. Mora (2006) extends this research by analyzing the east-Asian financial crisis and comes to the conclusion that CRAs are rather conservative in times of crisis. Contrary to this conclusion, Ferri et al. (1999), having analyzed the east-Asian crisis as well, find that CRAs behave in a

¹See White (2009)

procyclical fashion during times of financial distress.

So far there are only a few paper investigating the connection between sovereign CDS and sovereign ratings. Arezki et al. (2011) look at the impact of sovereign rating changes on several financial markets, including the CDS market. Ismailescu and Kazemi (2010) conduct an event study on the reaction of sovereign CDS on sovereign rating changes for emerging economies. Afonso et al. (2012) also conduct an event study using European CDS, bond yield, as well as rating data. They conclude that the effects between ratings and the other two series is significant, but also bi-directional.

When analyzing ratings, the estimation method usually employed is a binary or ordered choice regression with ratings as a dependent variable and variables explaining defaults as regressors. Typical examples for this are Gande and Parsley (2005), Mora (2006) or Freitag (2014). In order to tackle our research question, a similar methodology is needed. The difference between the papers mentioned and my approach is that they use macroeconomic variables that are published on a quarterly basis, while I make use of financial data which comes at a daily frequency. One could argue that ratings are also published at a daily frequency, since at every given trading day, a rating can be published, but this ignores the work that a CRA invests into a rating. Indeed, the typical rating process takes at least around a month for the big three agencies, as documented in Fuchs and Gehring (2013). This means that an estimation method is needed that is able to correlate data which are sampled at different frequencies. There are several possibilities to approach this issue. The first possibility is to simply average over the high frequency data. This is a convenient solution, however it causes an enormous loss of information contained in the high frequency series. Another possibility is a distributed lag model in which all current and past values (up to a certain point) are used as regressors to explain the low frequency variable. The problem with this approach is that it quickly leads to overfitting. The most recent innovation in the field of mixed frequency data was the introduction of MIDAS by Ghysels et al. (2004). The main innovation of MIDAS is that it lets the data determine its lag-length and weighting. This is done by using a polynomial with a small amount of parameters to weight the high frequency data. The value of these parameters are determined by optimizing them together with a slope coefficient using a minimum distance-type estimation. Therefore, MIDAS methods still use a substantial amount of information contained in high frequency data, while at the same time being a relatively parsimonious model to estimate.

Early examples of MIDAS application is Ghysels et al. (2005). It shows that the theoretical sound, but empirically rather elusive relationship between risk and return can be investigated in a time series setting by using MIDAS estimations. More recent applications of MIDAS are focused on forecasting of GDP or other economic indicators by using financial assets as regressors in addition to the more traditional explanatory variables. Typical examples are Clements and Galvão (2008) or Schumacher and Breitung (2008). Also, a wide array of different MIDAS derivatives were developed. Guérin and Marcellino (2013) construct a Markov-Switching MIDAS and apply it to US GDP-growth as well as US industrial production. A Factor MIDAS was developed by Marcellino and Schumacher (2010) who introduce three different MIDAS estimation methods and join them with three methods for factor analysis. Their methods are then applied to German GDP. Foroni et al. (2012) introduce an unrestricted MIDAS. This is a method of estimating a MIDAS with OLS, which works well if the difference in sampling frequency is relatively small. A smooth transition MIDAS is formulated by Galvão (2013) and applied to out-of-sample US and UK output growth. Foroni and Marcellino (2013) give a good overview of the current development in Mixed Data sampling. However, to the knowledge of the author, no binary choice MIDAS estimator has been developed, which would be the estimator of choice for this research. Therefore, in this paper a probit-MIDAS estimator is developed, as well as a procedure for testing the slope coefficient and an application to European rating data is done.

The rest of the paper is structured as follows. In section 2 CDS and ratings are discussed, in section 3 a short introduction to MIDAS-estimator is given, followed by the probit-MIDAS estimator. Section 4 contains simulations for the estimator, and in section 5 I estimate the relationship between sovereign CDS and sovereign rating changes. Section 6 concludes.

2 The Economics of CDS and Credit Ratings

In this section, the economics of CDS and the interaction with credit ratings are explained. CDS are a derivative that is insuring the buyer against the default of the asset underlying the CDS. The CDS works in the following way: The seller issues a contract promising to pay whoever is holding the contract, on default of the underlying asset, namely the nominal value of that asset. The buyer of a CDS is paying the seller a price which might be paid in one installment, in which case it is called the premium, or in staggered installments which are referred to as spread. The premium or spread can be calculated using the following formula:

$$\sum_{i=1}^N e^{-rt} Q(t_i) p = \int_0^{t_N} e^{-rt} (100 - M_q) q(t) dt, \quad (1)$$

where r is the risk free rate, $Q(t)$ is a risk neutral survival probability, M_t is the market value, and p is the CDS premium. The left-hand-side of this equation is the spread of the CDS, the right-hand-side is the insurance part which guarantees the reimbursement of any losses incurred after the underlying asset goes bankrupt. The only component that needs to be determined is $Q(t)$. This is done by matching the underlying asset with a risk-free counterpart of the asset. For example, the counterpart of an Italian government bond, would most likely be a German government bond with otherwise identical characteristics. The difference in price between these two bonds can be used to calculate the default risk that the market attaches to the non-risk-free bond. For further details, see Hull and White (2000).

In contrast to CDS which are a recently introduced derivative, Credit Rating Agencies have existed since the beginning of the 20th century. Starting in 1909, with the founding of Moody's, CRAs became important actors in financial markets after the great depression in 1929. During that time the ratings of CRAs became binding for a large number of market participants due to regulation by the Security Exchange Commission (SEC). These regulations also helped establishing the oligopoly that is currently present in the markets structure by making merely the ratings of Moody's, Fitch, Standard, and Poor's (back then still two separate entities) binding. In the seventies, the CRAs switched from a user-pays to an issuer pays system. This change came mainly because of the advent of copy machines which threatened the original business model of CRAs of selling booklets containing their ratings. From 1980 onwards, the CRAs became more involved into the business of sovereign ratings, mainly because developing countries were trying to access international capital market and needed a seal of approval for their bonds. In the last decade the rating agencies played a major role in the subprime debacle and the financial crisis by assigning top ratings to collateralized debt obligations (CDOs) consisting mainly of US mortgages. A lot of these mortgages were defaulting during the financial crisis despite the AAA rating given to the CDOs which contained them.²

CDS as well as credit ratings are measuring very similar things. While a rating is an estimate about the default probability of an entity, given by a CRA, in case of a CDS this default probability is reflected by its premium. However, the way ratings and CDS premiums are created is very different. A rating is constructed by a specialist at a CRA who takes into account macroeconomic factors as well as other

²For an overview of the history of credit rating agencies see White (2010)

country specific factors. CDS premiums are constructed using equation 1 and therefore the main driver are bond prices. This means that the premium of a CDS is driven by financial markets. This difference in construction might also dilute the correlation between these two variables. What follows is a discussion of possible reasons.

The first reason that needs to be mentioned is liquidity in the financial markets. Indeed, the lower the trading volume of a bond is, the more likely it is to have price distortion. The CDS premium is a nonlinear transformation of this price. Therefore, any reduced liquidity in the bond market directly affects the CDS premium. This would imply that CDS premium are increasing disproportionate in times of liquidity crunches. In terms of correlation this would imply that there is either a negative correlation or no correlation between CDS premium and ratings. However, this is not what we see in the majority of the results. The next issue that needs to be discussed is a possible decoupling of individual country characteristics from bond prices. Oliveira et al. (2012) provide some evidence that this is the case before 2007. Afmann and Boysen-Hogrefe (2012) shows that coefficients for macroeconomic factors in the Eurozone are time varying, with a noticeable increase after the crisis. However, the data used in this paper has its earliest point in 2006, and so this study uses mainly crisis or post-crisis data. Therefore, this issue should have no major impact in the estimations.

Looking again at the arguments, the above mentioned issues should not be a concern for the research question at hand.

3 Methodology

In this section the Probit-MIDAS estimator is described. It is an extension of the existing MIDAS estimator first introduced by Ghysels et al. (2004). The idea behind the MIDAS estimator is to use high frequency data, such as asset returns, to explain low frequency variables such as GDP or inflation rates. This is done by imposing a parametric weighting scheme on the high frequency variable(s). The main advantage of using MIDAS is that the weighting can extract information relevant for the estimation from extremely noisy data. Previously these questions were addressed using event studies such as in the previous mentioned paper of Afonso et al. (2012). The problem with this approach is that is difficult to make sure that only the event is actually affecting the asset prices. Additionally in case of an event study one has to rely on an asset-pricing model which introduces two more potential sources of errors, an estimation error as well as a specification error. The former is introduced by simply estimating the model. The latter comes from the fact that to date no general consensus on an universal asset pricing model has been reached.

3.1 The MIDAS estimator

I start by describing the MIDAS estimator. A simple MIDAS data generating process (DGP) can be written as:

$$y_t = x_t^{(m)}(\theta)\beta + u_t, \quad (2)$$

where:

$$x_t^{(m)}(\theta) = \sum_{j=1}^q w_j(\theta) L^{j/m} x_{t-j/m}^{(m)}, \quad (3)$$

with m being the sampling frequency of the high frequency variable, L is the lag-operator, and $w_j(\theta)$ is a polynomial function for weighting the regressor data. In order to estimate this Andreou et al. (2010)

show that only two assumptions are needed:

1) u_t is white noise.

2) $0 \leq w_j(\theta) \leq 1$ and $\sum_{j=1}^q w_j(\theta) = 1$.

This allows the identification of the slope coefficient. A widely used functional form for the weighting polynomial is the Almon lag polynomial:

$$w_j^a(\theta_1, \theta_2) = \frac{a(\frac{j}{j_m}, \theta_1, \theta_2)}{\sum_{j=1}^m a(\frac{j}{j_m}, \theta_1, \theta_2)}, \quad (4)$$

with

$$a(\frac{j}{j_m}, \theta_1, \theta_2) = \exp(\theta_1 j + \theta_2 j^2). \quad (5)$$

The estimations are executed using Nonlinear Least Squares (NLS). Andreou et al. (2010) show that the estimations are asymptotically efficient and in most cases superior to other approaches which are able to deal with data sampled at different frequencies such as averaging the high frequency data, or using distributed lag models. However, the main advantage of MIDAS compared to these other models is that on the one hand it is data-driven, that is the weighting scheme is estimated and therefore relatively little assumption on the exact weighting of the data is needed. On the other hand it is also a rather parsimonious approach which only needs to estimate one or two additional parameters. Ghysels et al. (2007) note that in principle non-linear MIDAS estimator are possible. Several examples are mentioned in the introduction of this paper such as the Markov-Switching MIDAS estimator by Guérin and Marcellino (2013).

3.2 The probit estimator

Binary choice methods in general, and the probit estimator in specific has long been a working horse of empirical microeconomics, although it has also seen some application in macroeconomics, for example Candelon et al. (2010). Probit estimations assume a latent variable DGP such that:

$$y_t^* = x_t \beta + \epsilon_t, \quad (6)$$

and

$$y_t = \begin{cases} 1 & \text{if } y_t^* > 0 \\ 0 & \text{if } y_t^* \leq 0, \end{cases} \quad (7)$$

In order to conduct estimations on this kind of data, a maximum likelihood approach is needed. The log-likelihood function(LLF) is:

$$l_t(\beta) = y_t \log[\Phi(x_t \beta)] + (1 - y_t) \log[1 - \Phi(x_t \beta)], \quad (8)$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution function. Apart from the standard maximum likelihood assumptions, the following additional assumptions are needed:

1) The error ϵ_t needs to be standard normal.

2) The data needs to follow a latent process from equation 7.

These can be found in Wooldridge (2001) and Greene (2008).

3.3 The probit-MIDAS estimator

Consider now a combination of the MIDAS DGP and the probit DGP in the form of

$$y_t^* = x_t(\theta)\beta + e_t, \quad (9)$$

where e_t has a normal distribution with mean zero and variance one. However, y_t is not actually observable but instead also governed by equation 7. In order to conduct estimations on this kind of data, one has to combine the probit and the MIDAS approach. The most straightforward way to do so is to embed the optimization routine for the MIDAS polynomial into the log likelihood function of a probit regression which yields the following log-likelihood function:

$$l_t(\beta, \theta) = y_t \log[\Phi(x_t^{(m)}(\theta)\beta)] + (1 - y_t) \log[1 - \Phi(x_t^{(m)}(\theta)\beta)]. \quad (10)$$

This estimation method needs to fulfill all assumptions for both approaches. This means the following assumptions need to be fulfilled:

- 1) The error e_t needs to be standard normal.

This is the standard error assumption for probit regressions. Note that it is placing additional restriction on the MIDAS assumption number one. Instead of requiring only finite variance, it is now restricted to be equal to one. Also, the distribution of the errors is now specified.

- 2) $0 \leq w_j(\theta) \leq 1$ and $\sum_{j=1}^q w_j(\theta) = 1$.

This is the same assumption as in the standard MIDAS model and the polynomials are always constructed in such a way that this assumption is fulfilled.

- 3) The data needs to follow a latent process described in equation 7.

Additionally the standard MLE assumptions from the textbook are needed.

3.4 Testing the $\beta(\theta) = 0$ Hypothesis

MIDAS regressions are often used for forecasting of various economic series, such as in Clements and Galvão (2008) or ?. But an economist might also be interested in the significance of the slope coefficient(s). However, MIDAS is problematic when one wants to test the significance of said coefficient. This is due to the non-identification of the weighting parameters under the null hypothesis. To understand the problem, let us assume that the slope coefficient of a MIDAS regression is actually equal to zero. This means that the weighting parameters are no longer uniquely identified. Looking at the alternative hypothesis it is obvious the weighting parameters can only take the value determined by our estimations, and are therefore uniquely identified. This means we have a lopsided scenario in which not all parameters are identified under one hypothesis but all of them are identified under the other hypothesis. Clearly, this is not a desirable scenario to conduct inference. Therefore, a recipe for the probit-MIDAS, which is bypassing this problem, is presented in this section. This involves a two-step procedure as well as a bootstrap. This approach is proposed by Ghysels et al. (2007). However, to the knowledge of the author this paper is one of the first to apply the bootstrap and provide simulation results for it.

Davies (1987) recognizes the problem of non-identified nuisance parameters under one type of hypothesis. He proposes to estimate over a grid of nuisance parameters and take the supremum instead of conducting standard inference. His paper provides asymptotic behavior for a simple test under these conditions.

Hansen (1996) takes this approach and tries to generalize it. He notes that the distributions of these tests depends upon a large number of parameters which makes deriving asymptotic behavior tedious as well as difficult to generalize. Instead he proposes a bootstrap approach. For the case of a probit-MIDAS this procedure can be implemented as follows:³

- 1) Estimate a probit-MIDAS regression using equation 10.
- 2) Construct a grid of c different θ combinations.
- 3) Use this grid to weight the regressor data, producing c different regressor series, and estimate c standard probit regressions.
- 4) Take the supremum over the t-statistics of the coefficient of these regressions as explained in Davies (1987).

Now we have a t-statistic. As a next step the critical value for this t-statistic needs to be computed. This is done by bootstrapping this procedure under the null hypothesis. What follows is the bootstrap loop:

- 5) Start by drawing bootstrap samples with replacement from the residuals.
- 6) Construct the dependent variable by adding up residuals and intercept.
- 7) Repeat step 3) and 4) from the first part with the bootstrap sample.
- 8) Repeat step 5) to 7) k times, where k is the amount of bootstrap samples.
- 9) Take the 95th percentile of the t-statistics to produce a critical value against which the t-statistic from step 4) can be evaluated to test the significance of the slope coefficient.

The author feels that one remark with respect to the actual bootstrapping is in order. A potential problem with the bootstrap is that there are no straightforwardly defined residuals in a probit regression. The standard solution in the literature to this issue is to use generalized residuals due to Gouriéroux et al. (1987), see for example Hsiao et al. (2012). However, one assumption when estimating a probit is that the errors are distributed as a standard normal random variable. Generalized residuals seem to consistently have a significant lower variance than one in this setup, and therefore seem to be unsuited for constructing the bootstrap sample. Possible other candidates are Pearson residuals, response residuals, and deviance residuals (see Hinkley et al. 1991). In section 4.2, a small-scale simulation study is done concerning the behavior of the different residual options available. In general, it is advised to inspect the residuals before starting the bootstrap.

4 Simulations

4.1 Finite Behavior of the Estimator

I start by investigating how the probit-MIDAS estimator is behaving in finite sample given different weighting schemes and different sampling frequencies. Then I investigate the behavior of the bootstrap. Starting with the first part, two different polynomials will be investigated. Apart from the Almon polynomial detailed in equation 5, I also include the Beta polynomial first introduced in Ghysels et al.

³This example uses only one regressor series

(2007) which is:

$$w_j^b(\theta_1, \theta_2) = \frac{b(\frac{j}{j^m}, \theta_1, \theta_2)}{\sum_{j=1}^m b(\frac{j}{j^m}, \theta_1, \theta_2)}, \quad (11)$$

with

$$b\left(\frac{j}{j^m}, \theta_1, \theta_2\right) = \frac{(\frac{j}{j^m})^{\theta_1-1} (1 - \frac{j}{j^m})^{\theta_2-1} \Gamma(\theta_1 + \theta_2)}{\Gamma(\theta_1) \Gamma(\theta_2)}, \quad (12)$$

where Γ represents the gamma-function. It is as flexible as the Almon polynomial and offers a wide variety of shapes. Both parameters need to be positive. Also both parameters are constrained from above (by 15), to allow for better convergence. The Almon polynomial parameters are constrained from above and below by -1 and 1, because otherwise the function becomes non-smooth. When comparing the two weighting schemes, the following Data Generating Process (DGP) is used:

$$y_t = \beta_1 x_t^{(m)}(\theta) + \beta_2 z_t + e_t. \quad (13)$$

Whereas $x_t^{(m)}(\theta)$ is a high frequency regressor, z_t is sampled at the speed as the dependent variable, and e_t is $N(0, 1)$. The reason for designing the DGP in such a way is that one can compare the estimations of the high frequency estimator with that of a low frequency regressor by comparing the performance of the first regressor to that of the second regressor. The simulations for every case investigates three different weighting schemes. These are downward sloping weights, upward sloping weights and a hump shape. The weighting in the DGP is always generated by the same polynomial that is used for the estimation. That is when using the Almon polynomial as a weighting polynomial in a probit-MIDAS estimation, the high frequency variable in the DGP is also generated by the Almon polynomial.

Typically MIDAS estimator are working with weekly data ($m = 5$), monthly data ($m = 22$), or quarterly data ($m = 66$). In this paper I only look at the first two, that is monthly frequencies, and weekly frequencies. The reason for this lies in the characteristics of the rating variable. In most cases, when considering a MIDAS approach, it is a priori clear in which sampling frequency the low frequency variable is sampled at. For example quarterly GDP estimates have a predefined date each quarter on which they are published. Ratings do not have a pre-specified date when they are announced. Thus, it is better to work with weekly or monthly regressor frequency. The intuition is that if one uses quarterly data, the ratings that are published in consecutive quarters could be up to 6 month apart. This seems to be an unacceptable difference between estimation and reality. Therefore, two different sampling frequencies are considered, a monthly one and a weekly one, since it could be argued that a month still gives too many different possibilities for a rating announcement, however, a week worth of high frequency data might simply not contain enough information for the estimations.

The first case to investigate is the monthly sampling frequency case, that is, $m = 22$. The DGP for the simulation is given by equation 13 with $\beta_1 = 0.5$ and $\beta_2 = 0.3$. The Almon polynomial parameters are (0.035, -0.085) for generating a downward sloping weighting, (-0.02, 0.005) for an upward sloping weighting, and (0.005 - 0.0005) for having a hump shaped weighting scheme. For the Beta polynomial I use (0.6, 2.1) to get a downward sloping weighting scheme, (1.1, 1) for generating an upward slope, and (4, 2.1) to have a hump shaped pattern. The parameters are chosen in such a way that they generate the appropriate shape in the given sampling frequency.⁴

All simulations are done with 2000 replications, the sample sizes are 50, 100, 250, 500, 750, and 1000. The tables are based on $\frac{\beta}{\beta^t}$ with β^t being the true parameter underlying the latent process. This should, due to the identification of the probit coefficient which is $\frac{\beta}{\sigma_e}$, and the fact that the errors are normally distributed

⁴Additionally there were some conflicts with the procedure supplying the initial values, and it was deemed safer to use parameter that are not exactly the same as those covered by this procedure.

with $\sigma = 1$, converge to 1. The deviations from this are denoted in percentages. The θ -parameter are not reported.

Starting with the Almon polynomial all three cases are displayed in table 1. In the first line of each case there is always the low frequency variable, which is the benchmark to which we compare the performance of the probit-MIDAS. As one can see the three cases are quite different from each other. The estimator performs reasonably well in a downward sloping weighting scheme, however the sample size to achieve full convergence seems quite high. When looking at the other two cases, that is upward sloping and hump shaped weighting scheme, one can see that the estimator does not perform well at all. In both cases there are still significant biases in the slope coefficients, even at a sample size of 1000. Fortunately these two cases are much rarer in economic applications than the downward sloping one.

Table 1: Almon polynomial simulations monthly frequency

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	-7.3 %	-15.6 %	-19.0 %	-6.2 %	-3.1 %	-1.6 %
	β_2	12.2 %	5.4 %	3.3 %	1.5 %	0.5 %	0.3 %
upward sloping weighting	$\beta_1(\theta)$	-22.2 %	-33.4 %	-39.0 %	-26.5 %	-18.9 %	14.1 %
	β_2	14.8 %	4.4 %	2.1 %	1.9 %	1.0 %	-0.0 %
hump-shaped weighting	$\beta_1(\theta)$	-49.5 %	-67.9 %	-77.0 %	-72.4 %	-65.8 %	-62.4 %
	β_2	14.4 %	6.4 %	2.4 %	0.9 %	0.1 %	0.2 %

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

The second polynomial that is investigated is the Beta polynomial. Starting again with the downward sloping case, one can see that the estimator is approaching the true value rather quickly. With a sample size of 250 the coefficient approaches very closely the true value and fluctuates around it when increasing the sample size. These results resemble the results of the Almon polynomials when it has a sample size of 1000 to estimate on. In line 3 and 4 of table 2 an upward sloping weighting scheme is investigated. The estimations are performing similar to the downward sloping weighting scheme. This is clearly superior to the downward sloping scheme using the Almon polynomial. As a last case a hump-shaped polynomial is investigated. Observing a very similar pattern as in the downward sloping case, the estimator achieves convergence to the true value relatively quickly and with a sample size of 250 the coefficient is already very close to the true parameter. However, there is also a disadvantage to this performance. The Beta polynomial has a relatively high number of cases (around 10 %) where the algorithm exits without finding a proper value for the slope coefficients and parameters.

Table 2: Beta polynomial simulations monthly frequency

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	11.2 %	0.4 %	2.2 %	2.2 %	4.4 %	0.7 %
	β_2	17.5 %	7.1 %	2.4 %	1.1 %	1.1 %	0.7 %
upward sloping weighting	$\beta_1(\theta)$	10.4 %	6.2 %	1.9 %	-3.1 %	-1.6 %	-1.8 %
	β_2	17.1 %	8.2 %	3.3 %	1.1 %	1.4 %	0.9 %
hump-shaped weighting	$\beta_1(\theta)$	11.2 %	-1.0 %	0.3 %	0.3 %	0.3 %	0.4 %
	β_2	12.9 %	6.9 %	2.5 %	0.8 %	0.7 %	0.4 %

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

As a next step a weekly sampling frequency is investigated, that is $m = 5$. The DGP is the same as in the monthly simulations. For the Almon polynomial I use $(-0.5, 0.085)$ to create a downward

sloping weighting scheme, $(0.5, 0.085)$ to build an upward sloping weighting scheme, and $(0.5, -0.085)$ to have an inverted-U shaped weighting scheme. For the Beta polynomial $(0.9, 1)$ is used for a downward sloping weighting, $(1.1, 1)$ is creating an upward sloping scheme, and $(4.1, 2.1)$ is for an inverted-U shaped weighting scheme. Other wise the setup is identical to the previous simulation case, 2000 replications are done and the sample sizes are 50, 100, 250, 500, 750, and 1000. The results for the Almon polynomial are in table 3

Table 3: Almon polynomial simulations weekly frequency

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	38.5%	16.6%	7.8%	4.8%	0.0%	0.0%
	β_2	13.3%	5.2%	2.5%	0.9%	3.8%	1.6%
upward sloping weighting	$\beta_1(\theta)$	-4.0%	-5.5%	0.9%	5.1%	4.8%	4.2%
	β_2	10.8%	4.9%	2.3%	1.0%	0.8%	0.5%
hump-shaped weighting	$\beta_1(\theta)$	19.6%	9.6%	4.5%	2.1%	2.7%	1.5%
	β_2	12.6%	5.9%	1.9%	0.8%	0.1%	0.5%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

The results for the Almon polynomial in this set of simulations looks good. The slope coefficient is converging for the downward sloping case, as the number of observation increases to no detectable bias at all. The distortion in the estimation when assessing the hump shaped pattern is relatively small. There is still some bias in the upward sloping weighting scheme, however, when comparing these results to the results obtained when using a monthly frequency, one can see a significant improvement over those results.

Table 4: Beta polynomial simulations weekly frequency

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	-15.1%	-21.6%	-24.3%	-23.2%	-20.1%	-18.7%
	β_2	14.3%	6.9%	1.8%	0.7%	0.5%	0.7%
upward sloping weighting	$\beta_1(\theta)$	-15.1%	-28.0%	-27.5%	-24.5%	-24.3%	-23.2%
	β_2	10.5%	2.0%	2.1%	0.7%	0.3%	0.1%
hump-shaped weighting	$\beta_1(\theta)$	8.8%	4.6%	1.1%	1.4%	1.3%	0.5%
	β_2	14.9%	6.6%	3.2%	0.1%	0.9%	0.7%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

The second set of simulations investigates the Beta polynomial. Also in this case does one see a significant difference compared to the previous set of simulations, however, here there is no improvement, but a worsening of the performance of the estimator. For the first two weighting schemes, there is a heavy bias, even when the sample size is at 1000. Only in the hump shaped case does one see an acceptable result for the estimator. Furthermore, compared to the monthly frequency case, the number of simulations where the algorithm exits the simulation without finding acceptable results is vastly increased, and in some cases up to 50 %.

Therefore, I come to the following conclusions. For the monthly sampling frequency, the best weighting scheme seems to be produced by the Beta polynomial. It needs significantly fewer observations in converging to the true parameter. Furthermore it does a good job in converging when dealing with a hump-shaped pattern. The Almon polynomial does not converge to its true parameter in this case. Contrary to the results for monthly frequency, for the weekly sampling frequency, the Almon polynomial seems to be doing the best job. Convergence to the true parameter is achieved with a modest sample size, and it can model a variety of shapes.

The obvious question is why does the Almon polynomial do such a bad job in the monthly setup, compared to the Beta polynomial and vice versa for the weekly setup. Starting with the monthly estimations, it should be noted that for the upward sloping weighting and the hump-shaped weighting, the Almon polynomial tends to prefer corner solutions. That is, at least one of the weighting parameter is equal to one. This in turn tends to lower the value of the regression coefficient, which explains the large negative deviations from the true coefficient value in table 1. The biggest difference between the two polynomials is a significantly increased forced exit rate when using the Beta polynomial in the monthly sampling frequency. A forced exit means that the optimizations algorithm is unable to converge in the given amount of iterations. With the Almon polynomial, this rate was around 0.1%, however using the beta polynomial this shot up to around 10 %. This means that in this case it is more obvious for the algorithm that the estimates are clearly off and it counteracts this by eliminating the faulty estimates. Therefore, the Beta polynomial does not suffer from the corner-solution case that can be found with the Almon polynomial. The forced exit rate is most likely also the reason for the difference in performance in the weekly sampling case. The Beta polynomial already has slight problems converging with significantly more observations, so it seems logical that for the weekly case the forced exit cases are increased. Indeed, the forced exit rate for the Beta polynomial shoots up to 50 %, which causes the drop in point-estimate performance.

One interesting fact worth mentioning is that the estimator seems to do reasonable well in small sample sizes, perform relatively bad in medium sized sample sizes, and better again in large sample sizes. However the authors has not come to a decent explanation as to why this is the case.

Sofar we have seen that the estimator has overall good finite sample properties, however the estimator has the tendency to select corner solutions. This is obviously a problem if it happens in an empirical application. What follows are approaches to circumvent this problem. If one looks again at the estimations, the downward-sloping weighting scheme seems to perform best. Therefore, it would be good if one could impose such a weighting on the estimator, in case of non-convergence. Indeed there is such a polynomial in use in the MIDAS-literature. Ghysels et al. (2009) introduce the hyperbolic polynomial:

$$w_j^h(\theta_1) = \frac{h(\frac{j}{j^m}, \theta_1)}{\sum_{j=1}^m b(\frac{j}{j^m}, \theta_1)}, \quad (14)$$

with $h(\frac{j}{j^m}, \theta_1)$ being:

$$h(\frac{j}{j^m}, \theta_1) = \frac{\Gamma(j + \theta_1)}{\Gamma(j + 1)\Gamma(\theta)}. \quad (15)$$

Only one parameters needs to be estimated. This parameter is constrained such that $0 < \theta < 0.5$ in order to guarantee stationarity (see Tanaka 1999). Since there are no different weighting pattern to be investigated, I simply look at three different parameter values. For the monthly frequency these are: (0.11), (0.251), and (0.41). The results are in table 5.

Table 5: Hyperbolic Polynomial Simulations Monthly Frequency

θ	Coefficients	50	100	250	500	750	1000
$\theta = 0.11$	$\beta_1(\theta)$	36.4 %	25.6 %	18.1 %	15.8 %	10.4 %	9.5 %
	β_2	7.5 %	3.5 %	1.8 %	1.0 %	0.6 %	0.6 %
$\theta = 0.251$	$\beta_1(\theta)$	19.6 %	14.6 %	9.8 %	8.7 %	6.2 %	5.6 %
	β_2	10 %	4.3 %	1.4 %	0.8 %	0.2 %	0.9 %
$\theta = 0.41$	$\beta_1(\theta)$	4.8 %	10 %	1.2 %	2.0 %	-0.7 %	-1.8 %
	β_2	1.28 %	3.6 %	1.5 %	0.5 %	0.6 %	0.0 %

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

As it can be seen the results for a rather low parameter coefficient are distorted even with 1000 observations. However, the higher the parameter-value gets, the better the estimates are. When inspecting closer the θ parameter, it turns out that also the hyperbolic polynomial tends to corner solution. However, having this polynomial at ones disposal gives another possible approach to the data.

Table 6: Hyperbolic Polynomial simulations weekly frequency

Weighting-scheme	Coefficients	50	100	250	500	750	1000
$\theta = 0.11$	$\beta_1(\theta)$	19.6%	10.5%	5.4%	3.4%	2.5%	2.4%
	β_2	7.0%	4.4%	1.3%	0.3%	0.0%	0.0%
$\theta = 0.3$	$\beta_1(\theta)$	11.9%	4.8%	1.4%	-1.5%	-1.4%	-1.9%
	β_2	10.6%	4.5%	1.1%	0.6%	1.1%	0.4%
$\theta = 0.49$	$\beta_1(\theta)$	15.5%	7.6%	2.3%	1.0%	1.5%	0.3%
	β_2	11.5%	4.7%	2.2%	0.9%	0.4%	0.0%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

The results for the weekly frequency are in table 6. They are very similar to the results for the Almon polynomial for this frequency (Note the slight change in θ to accommodate the change in frequency). The estimator improves over the monthly frequency and behaves nice in finite sample.

As a next step, I explore alternative optimization algorithms. All estimations so far were conducted using a combination of the BHHH algorithm and the Newton-Raphson algorithm, implemented in GAUSS 12, using the Constrained Maximum Likelihood package. Normally these methods are sufficient for solving a Maximum Likelihood problem. However, there are also possibilities to solve more troublesome optimization problems. What follows are simulations using an alternative method. Corana et al. (1987) introduced Simulated Annealing into statistics. It is a gradient free optimization approach particularly good at escaping local minima.⁵ What follows is a short description of how the algorithm works and simulation results. Simulated Annealing works in the following way:

- 1) Draw from a uniform distribution with endpoints -1 and 1, and scale it by a step size for all parameters to be estimated, to produce a step.
- 2) Accept or reject the steps by a probabilistic rule which is governed by a global variable called temperature (which is a positive number). The lower the temperature, the higher the probability that the optimal step gets accepted.
- 3) Do this N_s times.
- 4) Adjust the step size in such a way that about 50 % of the current moves would be accepted as moves next time.

⁵Also a Covariance-Evolutionary Algorithm was investigated as an alternative, however the results were similar to the standard algorithms

5) Decrease temperature, check for exit conditions, otherwise go back to 1).

Simulate Annealing is closely related to the Metropolis-Hasting Algorithm due to the accept-reject rule in step 2), which is a special case of Metropolis-Hasting. Next, simulation results are presented. The setup is the same as before. That is, all simulations are done with 2000 replications, the Almon as well as the Beta polynomial are investigated with parameters (0.035, −0.085) for generating a downward sloping weighting, (−0.02, 0.005) for an upward sloping weighting, and (0.005 − 0.0005) for a hump shaped weighting scheme. For the Beta polynomial (0.6, 2.1) is used to get a downward sloping weighting scheme, (1.1, 1) for generating an upward slope, and (4, 2.1) for a hump shaped pattern.

Table 7: Almon polynomial simulations monthly frequency with Simulated Annealing

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	28.5%	23.5%	16.0%	9.4%	6.1%	6.1%
	β_2	14.6%	6.7%	2.3%	0.6%	1.1%	0.3%
upward sloping weighting	$\beta_1(\theta)$	-26.5%	-19.6%	-18.9%	-16.2%	-13.2%	-13.2%
	β_2	13.7 %	6.5%	1.9%	0.6%	0.9%	0.3%
hump-shaped weighting	$\beta_1(\theta)$	-17.6%	-24.2%	-18.3%	-17.7%	-19.8%	-14.9%
	β_2	17.2%	5.9%	0.8%	1.5%	1.1%	0.2%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

As one can see, the results in Table 7 are very similar to the results obtained by standard algorithms earlier. The downward sloping weighting scheme behaves well, whereas the other two schemes pose a challenge to the estimator. However, one thing is remarkably different, the amount of corner solutions in the θ parameters seems to be lower compared to the standard optimization methods.

Table 8: Beta polynomial simulations monthly frequency with Simulated Annealing

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	2.3%	2.3%	-3.2%	-7.4%	-5.8%	-6.5%
	β_2	16.6%	7.2%	2.5%	0.1%	0.8%	0.8%
upward sloping weighting	$\beta_1(\theta)$	-9.9%	-22.9%	-25.2%	-24.7%	-22.1%	-21.3%
	β_2	15.2 %	3.4%	2.1%	1.1%	0.7%	-0.1%
hump-shaped weighting	$\beta_1(\theta)$	-38.4%	-36.5%	-35.1%	-28.1%	-27.7%	-13.4%
	β_2	15.2%	5.7%	2.2%	0.1%	0.7%	0.5%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

In contrast to the earlier results on the Beta polynomial, in this table we see a clear deterioration in the quality of the estimates. This happens because the rejection rate is now close to zero, compared to the case of standard optimization algorithms, where it was close to 10 %. This indeed confirms the assumption voiced earlier, that the rejection rate is driving the preciseness of the estimates for the Beta polynomial.

Table 9: Almon polynomial simulations weekly frequency with Simulated Annealing

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	20.9%	7.3%	4.3%	0.6%	1.3%	0.8%
	β_2	14.5%	3.9%	2.5%	1.9%	0.7%	0.7%
upward sloping weighting	$\beta_1(\theta)$	33.8%	17.4%	9.5%	5.0%	4.7%	3.8%
	β_2	10.1 %	5.1%	0.6%	1.4%	0.8%	0.8%
hump-shaped weighting	$\beta_1(\theta)$	17.6%	11.8%	4.0%	1.7%	1.4%	1.9%
	β_2	16.5%	6.0%	2.0%	1.4%	0.6%	0.9%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

Table 10: Beta polynomial simulations weekly frequency with Simulated Annealing

Weighting-scheme	Coefficients	50	100	250	500	750	1000
downward sloping weighting	$\beta_1(\theta)$	-11.0%	-21.3%	-22.9%	-20.2%	-19.9%	-19.6%
	β_2	11.9%	4.7%	2.2%	-0.2%	0.5%	0.2%
upward sloping weighting	$\beta_1(\theta)$	-14.7%	-26.7%	-25.3%	-23.6%	-23.4%	-22.6%
	β_2	14.3%	4.6%	2.4%	0.2%	0.4%	0.6%
hump-shaped weighting	$\beta_1(\theta)$	-3.8%	-4.8%	-1.4%	0.2%	1.7%	0.3%
	β_2	14.2%	7.2%	1.4%	0.9%	1.0%	0.1%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

In Tables 9 and 10, the results for weekly-frequency simulations can be found. Overall the results are remarkably similar to the ones using standard optimization tools. However, again fewer corner solutions are observed, which puts the Simulated Annealing method ahead. Next, the hyperbolic polynomial is investigated:

Table 11: Hyperbolic Polynomial Simulations Monthly Frequency with Simulated Annealing

θ	Coefficients	50	100	250	500	750	1000
$\theta = 0.11$	$\beta_1(\theta)$	36.3 %	25.9 %	15.3 %	13.0 %	10.4 %	7.1 %
	β_2	9.3 %	4.5 %	1.6 %	1.4 %	1.6 %	0.0 %
$\theta = 0.251$	$\beta_1(\theta)$	24.1 %	13.6 %	6.7 %	8.5 %	3.8 %	1.2 %
	β_2	8.2 %	6.5 %	1.9 %	0.6 %	0.8 %	0.8 %
$\theta = 0.41$	$\beta_1(\theta)$	20.7 %	7.5 %	-0.1 %	-0.1 %	-2.5 %	-1.4 %
	β_2	8.0 %	4.5 %	2.6 %	-0.1 %	1.0 %	0.5 %

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

Table 12: Hyperbolic Polynomial simulations weekly frequency

Weighting-scheme	Coefficients	50	100	250	500	750	1000
$\theta = 0.11$	$\beta_1(\theta)$	18.4%	12.0%	6.7%	3.9%	3.7%	2.8%
	β_2	7.6%	6.4%	1.4%	1.2%	-0.1%	-0.1%
$\theta = 0.3$	$\beta_1(\theta)$	16.5%	6.3%	1.9%	1.8%	0.0%	0.5%
	β_2	8.2%	4.6%	1.7%	0.7%	0.6%	0.4%
$\theta = 0.49$	$\beta_1(\theta)$	9.1.5%	2.9%	-1.3%	-0.7%	-2.5%	-1.0%
	β_2	12.4%	5.2%	1.5%	1.4%	0.1%	0.5%

Simulations are done with 2,000 replications. The table shows deviations from the true coefficient in percentages

When looking at the results for the hyperbolic polynomial, one can see marginal improvement over the Quasi-Newton methods used earlier. However, it seems generally advisable to use Simulated Annealing, instead of standard optimization routines when estimating a probit-MIDAS regression, because the

occurrence of corner solutions is reduced.

4.2 Finite Behavior of the Bootstrap

In this section the finite behavior of the bootstrap approach is investigated. Two issues are addressed. First of all, I investigate the bootstrap itself. Second of all I look at the residuals in the bootstrap procedure. The reason for this is that residuals in a probit are not as straightforwardly defined as they are in standard Ordinary Least Squares (OLS) regression. In the latter case it is simply $y_t - x_t\beta$. However, due to the nonlinear nature of the probit regression, there are several possibilities to define the residuals. The most popular definition is due to Gourieroux et al. (1987), who define the generalized residuals as:

$$u_t = \frac{\phi(x_t(\theta)\beta)}{\Phi(x_t(\theta)\beta)(1 - \Phi(x_t(\theta)\beta))}(y - \Phi(x_t(\theta)\beta)). \quad (16)$$

Also, there are three other residual versions for a probit. These were introduced in Hinkley et al. (1991). These are the response residuals defined as:

$$u_t = y_t - \Phi(x_t(\theta)\beta). \quad (17)$$

The Pearson residuals are defined as:

$$u_t = \frac{(y_t - \Phi(x_t(\theta)\beta))}{\sqrt{(\Phi(x_t(\theta)\beta)(1 - \Phi(x_t(\theta)\beta)))}}, \quad (18)$$

and the deviance residuals:

$$u_t = \begin{cases} \frac{\sqrt{(-2\log\Phi(x_t(\theta)))}}{\sqrt{(-2\log(1 - \Phi(x_t(\theta)))}} & \text{if } y_t = 0 \\ \sqrt{(-2\log(1 - \Phi(x_t(\theta)))}} & \text{if } y_t = 1. \end{cases} \quad (19)$$

In the probit case the econometrician has to work under the assumption that the errors are distributed as a standard normal variable. All the point estimates are based on this assumption. Therefore, one would also want to draw from a residual distribution which is close to the assumed distribution. Since there is an abundant choice of residual specification for this regression, one has to take a closer look at the properties of the different residuals. I am mainly interested in the variance of the different residual specifications. For this, three sets of simulations are conducted one each for one of the three main weighting scheme. The downward sloping weighting scheme is conducted with an Almon polynomial, the upward sloping polynomial and hump shaped one are estimated using the beta polynomial. I simply recycle the values for the weighting parameters from the simulations in section 4.1. The frequency is monthly, and there are 1000 observations and 250 replications. The results for the standard deviations of the different residuals can be found in table 13.

Table 13: Residual Standard Deviation

	Pearson residuals	Generalized Residuals	Response Residuals	Deviance Residuals
downward	1.001	0.659	0.366	0.133
upward	1.001	0.662	0.366	0.053
hump	1.001	0.662	0.366	0.059

Simulations are done with 250 replications and sample sizes 1,000, $m = 22$, and the Beta polynomial is used

As it can be seen, there are clear differences between the various residual types. Since one prefers the residuals to have a variance of one, it should be clear that the Pearson residuals are preferred over

the other types.⁶ However, it should be stressed again that these simulations in no way allow general conclusion about the optimal choice of the residual-specification for the bootstrap. Therefore, before actually conducting the bootstrap, the variance of the extracted residuals should be investigated.

As a next step the size and power of the bootstrap was investigated. For this investigations, the bootstrap was simulated 250 times with 1000 observations in each replication in a monthly frequency setting. The Almon polynomial was used. To make sure that the estimator is converging correctly, a downward sloping scheme was used to serve as parameters, thereby setting θ to $(0.035, -0.085)$. The first run analyzes the size of the test. Thus, the DGP is generated by white noise. In this setup it turns out that the test is moderately oversized, since for a 10 % confidence level, the null hypothesis that the coefficient is equal to zero is 34 times (out of a total of 250) rejected. The second run investigates the power of the bootstrap. The data is determined by the following DGP:

$$y_t = \beta_1 x_t^{(m)}(\theta) + e_t, \quad (20)$$

with $\beta_1 = 0.5$. Again, 250 bootstraps were simulated. In this case the results are not looking as good as in the size test. It turns out that in 250 replications, we see a Type II error in 130 of the cases, given an $\alpha = 0.1$ the β (the probability of a Type II error) is around 0.52. This seems to be a rather low power and generally means that one should be suspicious when not being able to reject the null hypothesis. Indeed, when comparing the setup to a standard probit with the same slope coefficient, the β is close to 0. Therefore, the test is definitely not performing well on the power side.

A last remark in this section shall be devoted to the grid. So far no systematic investigation has been done to determine properties and optimal composition of the grid. Here, a brute force approach is used. Both the Almon polynomial as well as the hyper polynomial are bounded from above and below. Thus it is possible to simply create a monotonically increasing series between the two bounds, and in case that more than one weighting parameter is estimated, to compute the Cartesian product between them to construct a grid. For the beta polynomial, which is not bounded from above, one simply has to set a stop at the increasing series at some point. The experience so far with the construction of the grid reveals that a grid with number of combinations in the low double digit (10-20) is too small, but a grid in the low triple digit (100-150) is giving acceptable results. A denser grid (400) does not yield superior results.

5 Estimations

In this section the relationship between sovereign CDS data and sovereign rating will be analyzed. Specifically we want to know whether CRAs and CDS agree upon the default probability of sovereign nations. This is done by first regressing lagged differenced CDS premiums on downgrades using the probit MIDAS estimator, and next by doing an identical probit-MIDAS regression, but with lead CDS premiums instead of lagged ones. By doing so we can look at anticipation as well as market reaction.

Three European datasets are analyzed. First, I investigate 10-year CDS data. The availability of the series differs drastically amongst countries, therefore to maximize the amount of observations, two panels were created. One which is named the western European panel, encompassing Austria, Belgium, Cyprus, Finland, France, Germany, Greece, Ireland, Italy, Portugal, and Spain, which spans June 2009 till December 2012. A second panel named the eastern European panel has Bulgaria, Czech republic, Estonia, Hungary, Lithuania, Poland, Romania, Slovakia, and Slovenia in it, and covers April 2006 till September

⁶Exploratory simulations show that the very same problem is present in standard probit regressions

2009. Thirdly, I have a dataset of 5-year maturity CDS, consisting of Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Poland, Portugal, Slovenia, and Spain. The panel starts in June 2009 and ends in June 2013. All data has been taken from Datastream. It should be noted that the CDS premiums are not trading data produced by financial markets, but the premiums that buyers of CDS pay to acquire a newly issued CDS which are calculated as explained in section 2.

I start with a probit-MIDAS regression with downgrades of the country as the dependent variable, sampled at monthly frequency, and the lagged differenced⁷ sovereign CDS of the respective nation as an independent variable, such that we have:

$$downgrade_{i,t} = \beta_0 + \beta_1(\theta)\Delta CDSpremium_{i,t-1}^{(m)} + u_{i,t}. \quad (21)$$

If β is significant this shows that CDS have a similar opinion on default probabilities as CRAs.

A panel pooled approach is employed. This has the advantage that we have more information at hand to estimate the relationship and the disadvantage that it is assumed that the CDS markets for all countries are reacting in the same way to rating changes. The analysis starts with the monthly CDS data. Every rating movement that happened from the 20th of each month is instead assigned to the following month. The significance of the slope coefficient is tested with the bootstrap described earlier. However, instead of sampling with replacement over all residuals, sampling takes only place over the cross section, following Hansen (1999). For the estimations where $m = 22$, the hyper polynomial is used since the Almon polynomial tends to select corner solutions in a lot of the estimations conducted. The optimization method is Simulated Annealing, due to its slightly better performance. The results can be found in Table 14.

Table 14: Monthly Estimations

S & P	West Panel	East Panel	5-year maturity
intercept	-1.523	-2.198	-1.745
CDS	0.007**	0.061	0.060**
θ	0.436	0.002	0.498
Fitch			
intercept	-1.586	-2.256	-1.729
CDS	0.007**	0.060	-0.024**
θ	0.461	0.001	0.420
Moody's			
intercept	-1.521	-2.274	-1.746
CDS	0.006**	-0.076**	0.049**
θ	0.474	0.497	0.003

In the first column the analysis for the western panel can be found. For all three CRAs do we see positive significant results as expected by theory. Therefore, CDS spreads are increasing prior to a downgrade. The second column contains the eastern European panel. Here, we see a different pattern. Only for one agency is the coefficient significant. In the other two regressions the coefficients are statistically not different from zero. This implies two possibilities. First, financial markets are actually surprised by the downgrade and are reacting afterwards to it, or two, financial markets do not agree with CRAs on the default probability of a country and therefore ignore their decisions. To test this, another set of probit-MIDAS regression is estimated using lead instead of lagged CDS data. If the first explanation is true we should see a positive significant coefficient, if the second one is true we should see a non-significant or

⁷Since downgrades are essentially differenced ratings, the CDS data also needs to be differenced

negative coefficient. For the 5-year maturity data, we see that in two cases there is the expected positive significant case, and in one case the coefficient is negative. Again, the regressions are re-estimated using lead CDS data.

Next, the lead estimations are conducted. The results can be found in Table 15. Note that all regressions are re-estimated. The reason for this is that even in cases with a positive significant coefficient, there might still be some reaction after the downgrade, that is the downgrade is expected with a certain probability, but financial markets are not completely sure about the actions of the CRAs.

Table 15: Monthly Lead Estimation

S & P	West Panel	East Panel	5-year maturity
intercept	-1.504	-2.161	-1.698
CDS	0.008	-0.093***	-0.020*
θ	0.004	0.493	0.490
Fitch			
intercept	-1.569	-2.131	-1.710
CDS	0.002	-0.059	0.016
θ	0.399	0.484	0.023
Moody's			
intercept	-1.518	-2.242	-1.685
CDS	0.007**	-0.070**	-0.013
θ	0.020	0.484	0.035

We can see that in nearly all cases there are no reactions after the downgrade, indicated by a non-significant point estimate, or a decrease in CDS premiums. Only the estimations for Moody's with the west panel indicate a further premium increase for CDS. Therefore, it can be concluded that on a monthly level, the decisions of CRAs are expected by markets, for the cases where we see a positive significant coefficient for the lagged estimations. It should be additionally noted that the cases where we had significant negative coefficients in the lagged estimations, don't exhibit positive significant coefficients for the lead data case. Thus, there seems a difference in opinion on the default probability between CRAs and financial markets.

As a next step all regressions are re-estimated with rating changes sampled on a weekly frequency, therefore $m = 5$. The reason for doing this is that months do not have exactly 22 trading days as assumed by the framework for the monthly estimations. Instead in this sample the average month has 21.75 trading days. This means that as t increases, the lagged independent variable is slowly moving towards the same t as the dependent variable, which might lead to an endogeneity problem. Fortunately in our data-set, every week has exactly five days which circumvents this problem. Additionally, we can now look at the impacts on CDS premiums closer to the downgrade, which is especially interesting in the lead-estimations. Therefore, all regressions are again estimated with weekly frequency sample for the ratings. The western European panel now spans from the 15th of May 2009 until December 21st 2012, the eastern Panel covers 17th of March 2006 to 24th of September 2010, and the five-year maturity CDS panel goes from June 2009 till June 14th 2013. The results are in table 16.

Table 16: Weekly Estimations

S & P	West Panel	East Panel	5-year Maturity
intercept	-2.179	-2.726	-2.367
CDS	0.003**	0.030**	0.018*
θ_1	0.966	0.823	-0.31
θ_2	-0.197	-0.176	-0.050
Fitch			
intercept	-2.246	-2.785	-2.498
CDS	0.002	-0.021**	-0.005
θ_1	0.193	0.919	0.531
θ_2	-0.001	-0.573	0.008
Moody's			
intercept	-2.116	3.095	-2.353
CDS	0.002**	-0.059**	-0.012
θ_1	-0.042	0.965	0.568
θ_2	-0.015	-0.154	-0.167

The weekly estimation For the west panel paint a similar picture as the monthly estimations. In two cases we see positive significant coefficients, in one case is the coefficient insignificant, so in one case a positive significant coefficient was changed to a non-significant coefficient. This does not change the interpretation made for the monthly estimations.

For the east panel, we see that the coefficient for Fitch are switched from a negative to a positive one. Therefore, the decisions of this agency are not expected for the monthly horizon, however when we look at the weekly data, markets are expecting the downgrades simply within a smaller time frame. For the other two cases, there are significant negative coefficients which indicate that there is a disagreement between CRAs and CDS implied default probabilities.

The five-year maturity data shows in all three cases agreement with the monthly estimations. As a next step, all regressions are again re-estimated using lead CDS data. The results can be found in table 17.

Table 17: Weekly Lead Estimations

S & P	West Panel	East Panel	5-year Maturity
intercept	-2.153	-2.740	-2.350
CDS	0.000	-0.030**	-0.010*
θ_1	-0.285	0.947	-0.303
θ_2	-0.49	-0.441	-0.882
Fitch			
intercept	-2.251	-2.898	-2.528
CDS	-0.003*	0.058	-0.006**
θ_1	0.396	0.977	-0.555
θ_2	-0.042	-0.227	-0.854
Moody's			
intercept	-2.130	-2.830	-2.351
CDS	0.000	-0.026**	-0.014**
θ_1	0.366	0.896	0.918
θ_2	0.049	-0.017	-0.732

For the west panel lead estimation, we see that they are more or less in line with the monthly estimations. In one case do we see a decrease of CDS premiums, where we saw no significant movement for the monthly estimations. In another case there was a positive significant coefficient in the monthly estimations, where the corresponding weekly coefficient is insignificant. For the east panel we see exactly the same coefficients

as in the monthly lead regressions. The third column has three negative significant coefficients, whereas beforehand we saw that in only one case for the monthly estimations.

It should be noted that even with these deviations, generally the lead estimations for both frequencies have rather small coefficients even if they are significant and overall have the same tendency. Markets are calming down after a downgrade. Therefore, the lead estimations do not give rise to the theory that markets are surprised by downgrades and are reacting afterwards. Instead we have two cases. There is the evaluation of eastern European countries, where CRA decisions and implied default probability by CDS are clearly diverging. In the other cases, such as the western panel, ratings and CDS premiums are moving into the same direction and therefore financial markets agree with CRAs about default probabilities of sovereigns.

Also, I am doing the same regression but pooling all the agencies into one dependent variable.⁸ For the monthly series we see that in no case is the coefficient significant. Thus, the markets seem to look at individual CRA decisions rather than combining them together into one information set.

As a further test I look at the rating watch status of countries. Having a rating on watch means that it is due for a re-evaluation. Furthermore the CRAs mostly attach a tendency to this watch as either negative or positive. Therefore, it could be the case that CDS premiums do not react to downgrades, but to the announcement of a rating coming under scrutiny. Therefore all regressions are re-estimated using negative rating watch as a dependent variable. For the monthly case there are only three cases of significant coefficients. However, two of them are from the east panel which indicates that for these countries the rating-watch announcement are more interesting.⁹ The weekly regressions have two significant regression coefficient for the 5-year maturity data, but otherwise everything else is statistically not significant. This further supports the hypothesis that rating watch status is not important for the CDS market.

The interesting question is, whether one even needs CRAs, when it seems that CDS premiums can yield similar information about default probability already before the rating is released. However, for this claim one should keep in mind that CDS premiums are rather volatile, and that the employed MIDAS scheme is reducing the noise in the data significantly before correlating this data with downgrades. Therefore, it is unlikely to extract these information by simply looking at CDS premiums, and the CRAs are thus still valuable to the market.

6 Conclusion

In this paper it is investigated whether it is possible to forecast rating changes with the help of CDS premiums. This should give an indication of whether CRAs supply financial markets with new information, or whether investors already price this in beforehand. Also it is investigated whether downgrades and CDS premium are actually correlated with each other as expected by theory. To do so a probit-MIDAS estimator is developed to account for the characteristics of CDS time series data, which is available at rather high frequency on the one hand, and the event-type nature of rating changes on the other hand. While probit regressions are a long time workhorse of modern econometrics, MIDAS is a rather new type of estimator introduced by Ghysels et al. (2004). The idea behind a MIDAS estimator is to weigh the data using a parsimonious weighting scheme in which the parameters of the weighting functions are minimized jointly with the regression coefficient. The estimator is investigated using Monte Carlo simulations. Also, multiple weighting functions are explored. It turns out that the probit-MIDAS estimator

⁸Results are available upon request.

⁹Fitch was excluded due to no published negative rating watch, the results are available upon request.

performs differently for different weighting schemes and sampling speeds, but if the optimal weighting scheme is employed for a given sampling frequency, it behaves well in finite sample. Also, I want to test the significance of the slope-coefficient. This paper is one of the first to implement and investigate a bootstrap approach proposed by Ghysels et al. (2007). The test has reasonable size but is lacking power.

When investigating the relationship between sovereign CDS and sovereign downgrades, this paper uses various European countries with data stretching from 2006 until 2013. 10-year-maturity contracts as well as 5-year-maturity contracts are investigated. It turns out that for a monthly sampling frequency, CDS premiums and ratings do exhibit a positive significant relationship in numerous of the investigated cases. Multiple robustness checks, such as changing sampling frequency or exploring credit rating announcements confirm this. Therefore, CRAs and CDS data agree in a substantial amount of cases on the default probability of sovereign nations. Also, it should be noted that CDS premiums are expecting downgrades in advance, that is there is a positive significant price increase in a substantial amount of cases. However, this does not mean that raw CDS premiums can give a direct indication of the default probability of the underlying asset. By using a MIDAS estimator, I remove most of the noise from the data and this allows the extraction of default indication from CDS data. Therefore, rating agencies are still giving a valuable service to financial markets.

References

- Afonso, A. (2003). Understanding the determinants of sovereign debt ratings: Evidence for the two leading agencies. *Journal of Economics and Finance*, 27(1):56–74.
- Afonso, A., Furceri, D., and Gomes, P. (2012). Sovereign credit ratings and financial markets linkages: application to european data. *Journal of International Money and Finance*, 31(3):606–638.
- Alter, A. and Schüler, Y. S. (2012). Credit spread interdependencies of european states and banks during the financial crisis. *Journal of Banking & Finance*, 36(12):3444–3468.
- Andreou, E., Ghysels, E., and Kourtellis, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, 158(2):246–261.
- Ang, A. and Longstaff, F. A. (2013). Systemic sovereign credit risk: Lessons from the us and europe. *Journal of Monetary Economics*, 60(5):493–510.
- Arezki, R., Candelon, B., and Sy, A. (2011). Sovereign rating news and financial markets spillovers: Evidence from the european debt crisis. *IMF working papers*, pages 1–27.
- Arghyrou, M. G. and Kontonikas, A. (2012). The emu sovereign-debt crisis: Fundamentals, expectations and contagion. *Journal of International Financial Markets, Institutions and Money*, 22(4):658–677.
- Aßmann, C. and Boysen-Hogrefe, J. (2012). Determinants of government bond spreads in the euro area: in good times as in bad. *Empirica*, 39(3):341–356.
- Candelon, B., Dumitrescu, E., and Hurlin, C. (2010). Currency crises early warning systems: Why they should be dynamic. *METEOR research memorandum*.
- Cantor, R. and Packer, F. (1996). Determinants and impact of sovereign credit ratings. *Economic Policy Review*, 2(2):37–53.
- Clements, M. P. and Galvão, A. B. (2008). Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the united states. *Journal of Business & Economic Statistics*, 26(4):546–554.
- Corana, A., Marchesi, M., Martini, C., and Ridella, S. (1987). Minimizing multimodal functions of continuous variables with the “simulated annealing” algorithm. *ACM Transactions on Mathematical Software (TOMS)*, 13(3):262–280.
- Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 74(1):33–43.
- Ferri, G., Liu, L., and Stiglitz, J. (1999). The procyclical role of rating agencies: Evidence from the east asian crisis. *Economic Notes*, 28(3):335–355.
- Froni, C. and Marcellino, M. (2013). A survey of econometric methods for mixed-frequency data.
- Froni, C., Marcellino, M., and Schumacher, C. (2012). U-midas: Midas regressions with unrestricted lag polynomials. *CEPR Discussion Paper No. DP8828*.
- Freitag, L. (2014). Procyclicality and path dependence of sovereign credit ratings the example of europe. *GSBE research memorandum*.
- Fuchs, A. and Gehring, K. (2013). The home bias in sovereign ratings.

- Galvão, A. B. (2013). Changes in predictive ability with mixed frequency data. *International Journal of Forecasting*, 29(3):395–410.
- Gande, A. and Parsley, D. (2005). News spillovers in the sovereign debt market. *Journal of Financial Economics*, 75(3):691–734.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The midas touch: Mixed data sampling regression models.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509–548.
- Ghysels, E., Sinko, A., and Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90.
- Ghysels, E., Valkanov, R., and Serrano, A. (2009). Multi-period forecasts of volatility: Direct, iterated, and mixed-data approaches. In *EFA 2009 Bergen Meetings Paper*.
- Gourieroux, C., Monfort, A., Renault, E., and Trognon, A. (1987). Generalised residuals. *Journal of Econometrics*, 34(1-2):5–32.
- Greene, W. H. (2008). *Econometric Analysis*, 6/e. Pearson Education.
- Guérin, P. and Marcellino, M. (2013). Markov-switching midas models. *Journal of Business & Economic Statistics*, 31(1):45–56.
- Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica: Journal of the Econometric Society*, pages 413–430.
- Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of econometrics*, 93(2):345–368.
- Hinkley, D. V., Reid, N., and Snell, E. (1991). *Statistical theory and modelling: in honour of Sir David Cox, FRS*. Chapman and Hall.
- Hsiao, C., Pesaran, M., and Pick, A. (2012). Diagnostic tests of cross-section independence for limited dependent variable panel data models*. *Oxford Bulletin of Economics and Statistics*.
- Hull, J. C. and White, A. (2000). Valuing credit default swaps 1: No counterparty default risk. *NYU Working Paper No. FIN-00-021*.
- Ismailescu, I. and Kazemi, H. (2010). The reaction of emerging market credit default swap spreads to sovereign credit rating changes. *Journal of Banking & Finance*, 34(12):2861–2873.
- Marcellino, M. and Schumacher, C. (2010). Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp. *Oxford Bulletin of Economics and Statistics*, 72(4):518–550.
- Mora, N. (2006). Sovereign credit ratings: guilty beyond reasonable doubt? *Journal of Banking & Finance*, 30(7):2041–2062.
- Oliveira, L., Curto, J. D., and Nunes, J. P. (2012). The determinants of sovereign credit spread changes in the euro-zone. *Journal of International Financial Markets, Institutions and Money*, 22(2):278–304.
- Pan, J. and Singleton, K. J. (2008). Default and recovery implicit in the term structure of sovereign cds spreads. *The Journal of Finance*, 63(5):2345–2384.
- Remolona, E. M., Scatigna, M., and Wu, E. (2008). The dynamic pricing of sovereign risk in emerging markets: Fundamentals and risk aversion. *The Journal of Fixed Income*, 17(4):57–71.

- Schumacher, C. and Breitung, J. (2008). Real-time forecasting of german gdp based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, 24(3):386–398.
- Tanaka, K. (1999). The nonstationary fractional unit root. *Econometric theory*, 15(04):549–582.
- White, L. (2009). The credit-rating agencies and the subprime debacle. *Critical Review*, 21(2):389–399.
- White, L. (2010). Markets: The credit rating agencies. *The Journal of Economic Perspectives*, 24(2):211–226.
- Wooldridge, J. (2001). *Econometric analysis of cross section and panel data*. MIT press.